8. Searching Financial Patterns with Self-organizing Maps

Shu-Heng Chen[1] and Hongxing He[2]

[1] AI-ECON Research Group, Department of Economics, National Chengchi University, Taipei, Taiwan 11623
email: chchen@nccu.edu.tw
[2] CSIRO Mathematical and Information Sciences, Commonwealth Scientific & Industrial Research Organization, GPO Box 664, Canberra ACT 2601, Australia
email: Hongxing.He@csiro.au

Using *Self-organizing maps* (**SOM**), this paper formalizes chartists' behavior in searching of patterns (charts). By applying a 6 by 6 two-dimensional SOM to a time series data of TAIEX (Taiwan Stock Index), 36 patterns are established. To see whether these 36 patterns transmit profitable signals, a "normalized" equity curve is drawn for each pattern up to 20 days after observing the pattern. Many of these equity curves are either monotonically increasing or decreasing, and none of them exhibits random fluctuation. Therefore, it is concluded that the patterns established by SOM can help us foresee the movement of stock index in the near future. We further test profitability performance by trading on these SOM-induced financial patterns. The equity-curve results show that SOM-induced trading strategy is able to outperform the buy-and-hold strategy in quite a significant period of time.

8.1 Motivation and Introduction

It has been a long time that technical analysts detect trading signals with charts,[1] and for this, they are frequently called *chartists*. Since chartists has been a well-established profession in the financial industry, there is little doubt that charts, to some extent, did transmit signals. Nonetheless, from a scientific viewpoint, charts are somewhat subjective objects. In general, there is no well-articulated definition for charts. Alternatively speaking, analysts have to rely on their *experience* to identify charts. In this research, we attempt to provide a systematic study of charts. The ideas is simple: *if charts do transmit information, then the time series of assets' prices should have many unique patterns.* What we need is a mechanical tool to discover these patterns, and Self-organizing maps (**SOM**) seem to be a very natural tool to serve this purpose. In this paper, we use **SOM** to search for charts, and test whether these charts are profitable.

This paper is organized as follows. Section 2 is a brief introduction to SOM, and gives the experimental design given in this paper. Section 3

[1] Some of them even consider what they are doing as a science. See [8.2].

S.-H. Chen et al. (eds.), *Computational Intelligence in Economics and Finance*

presents the 36 patterns established by SOM. Section 4 investigates the information contents of these 36 charts. Trading performance of these 36 patterns are analyzed in Section 5 followed by a few concluding remarks in Section 6.

8.2 Self-organizing Maps

In contrast to the artificial neural networks which are used for *supervised learning*, there is another special class of artificial neural networks known as *Self-organizing feature map* (**SOFM** or, simply, **SOM**). The SOM is used for *unsupervised learning* to achieve *auto classification*, *data segmentation* or *vector quantisation*. The SOM adopts a so-called *competitive learning* among all neurons. The output neurons that win the competition are called *winner-takes-all* neurons. In SOM, the neurons are placed on the sites of a n-dimensional lattice. The value of n is usually 1 or 2. Through competitive learning, the neurons are tuned to represent a group of input vectors in an organized manner. The mapping from continuous space to a discrete one or two-dimensional space achieved by SOM reserves the spatial order.

Among a number of training algorithms for SOM, the Kohonen's learning algorithm is the most popular one ([8.5],[8.3]). Kohonen's learning algorithm adopts a heuristic approach. Each neuron on the lattice has a weight vector of m components attached. The m is the number of input variables of the input data sets. The winning neuron and its close neighbors in the lattice have their weight vectors adjusted towards the input pattern presented on each iteration. Unlike other clustering methods such as k-means clustering([8.4]), the Kohonen's SOM has an advantage that the final training outcome is insensitive to the initial settings of weights. Therefore, Kohonen's SOM has found a wide variety of applications.

In this paper we present the result of the application of the SOM to financial time series data. The data set to be segmented is the daily closing price of Taiwan stock indices (TAIEX). The original data set covers the daily price from 1/5/71 to 3/26/97, which has 7435 observations. A sliding window whose width covers 125 trading days moves from the first day to the 7309th day. This gives us 7309 subsamples (windows) each with 125 observations. Each subsample represents a time series pattern. The SOM is then used to automatically divide all patterns to groups or clusters in a way such that members of the same group are *similar* (*close*) in the Euclidean metric space. The 125 observations of each subsample is normalized between 0 to 1. A two-dimensional 6×6 SOM is used to map the 7309 records into 36 clusters.[2] The 6×6 lattice of SOM is presented in Fig. 8.1.

[2] Regarding the question why the two-dimension lattice model is used for SOM, we do not have a theoretical answer here. However, the following observations may lend support for the usage of a two-dimensional map. First, in practice, only low-dimension lattice (dimension no more than three) has been used. Second, in

0	1	2	3	4	5
6	7	8	9	10	11
12	13	14	15	16	17
18	19	20	21	22	23
24	25	26	27	28	29
30	31	32	33	34	35

Fig. 8.1. Pattern indices and their corresponding positions in 6×6 lattice of SOM

In the training process, the weights of the winning neuron and its close neighbors are updated according to Eq. (8.1),

$$w_j(n+1) = w_j(n) + \eta(n)\pi_{j,i(x)}(n)[x - w_j(n)], \tag{8.1}$$

where $w_j(n)$ is the weight vector of the jth neuron at the nth iteration, $\pi_{j,i(x)}(n)$ is the *neighborhood function* (to be defined below) of node indices j and $i(x)$,

$$i(x) = arg \min_j \| x - w_j \|, j = 1, 2, ..., N, \tag{8.2}$$

and $\eta(n)$ is the *learning rate* at iteration n.

We take for the neighborhood function the *Gaussian* form,

$$\pi_{j,i(x)} = \exp^{-\frac{d^2_{j,i(x)}}{2\sigma^2(n)}}, \tag{8.3}$$

where $\sigma(n)$ is some suitably chosen, monotonically decreasing function of iteration times n. Here, the effective width σ decays with n according to Eq. 8.4.

$$\sigma(n) = \sigma_0 \exp^{-\frac{n}{\tau_1}}, \tag{8.4}$$

where σ_0 and τ_1 are constants. The learning rate decays in a similar manner:

$$\eta(n) = \eta_0 \exp^{-\frac{n}{\tau_2}}, \tag{8.5}$$

where η_0 and τ_2 are constants.

The training takes a long time with initially almost all neurons have their weights updated. This training phase is called the *ordering phase*. The

---

the commercial package, only low-dimension maps are implemented, and almost all use the two-dimensional lattice as a default setting. Third, in his book ([8.6]),

Table 8.1. Parameter setup for the implementation of the 2-dimensional 6×6 SOM

Number of total trading days	7435
Number of time series pattern	7309
Number of trading day in each time series	125
Dimensionality of SOM	2
Number of neuron	36
Ordering phase initial learning rate	0.900
Ordering phase learning rate decay rate	455.120
Ordering epoch	1000
Ordering phase initial radius	8.485
Ordering phase radius decay rate	467.654
Convergence phase initial learning rate	0.100
Convergence learning rate decay rate	333.808
Convergence phase initial radius	1.000
Convergence phase radius decay rate	434.294
Convergence phase epoch	1000

weights then are settle down gradually at the second phase of learning named *convergence phase* where only the weights of winning neuron and perhaps its nearest neighbors are updated according to the case presented. The control parameters chosen to conduct this experiment is summarized in Table 8.1.

8.3 Charts Constructed by SOM

After the training, each of the 7309 time series subamples (charts) is assigned a cluster index which is represented by its winning neuron. The charts assigned to the same cluster index are *closer* than the charts which belong to difference clusters in terms of Euclidean distance. Then the average behavior of each cluster can serve as a *representative chart* for that cluster. Since we are using 6×6 SOM, in the end, we have 36 representative charts (See Figs. 8.2 and 8.3). These 36 charts can then be considered as something "*equivalent*" to what chartists are looking for.

How can these SOM-induced charts be compared to those charts used by chartists? Well, we have to admit that there is no direct comparison, knowing that each of our charts is *only* a *representative* (*average*) of many similar charts, and hence is not referred to any specific chart, while the charts used and defined by chartists are based on a single price line. Therefore, not

Kohonen mentioned that the original idea of SOM is enlighten by the structure of the human brain. The brain has a surface area of 2400 cm^2. The cerebral cortex is like a two-dimension map. It is divided into a few areas like motor cortex, somatosensory cortex, visual cortex, auditory cortex, etc. Each area is responsible for one kind of memory or control. Therefore, the two-dimension lattice model is closest to the structure of the human brain.

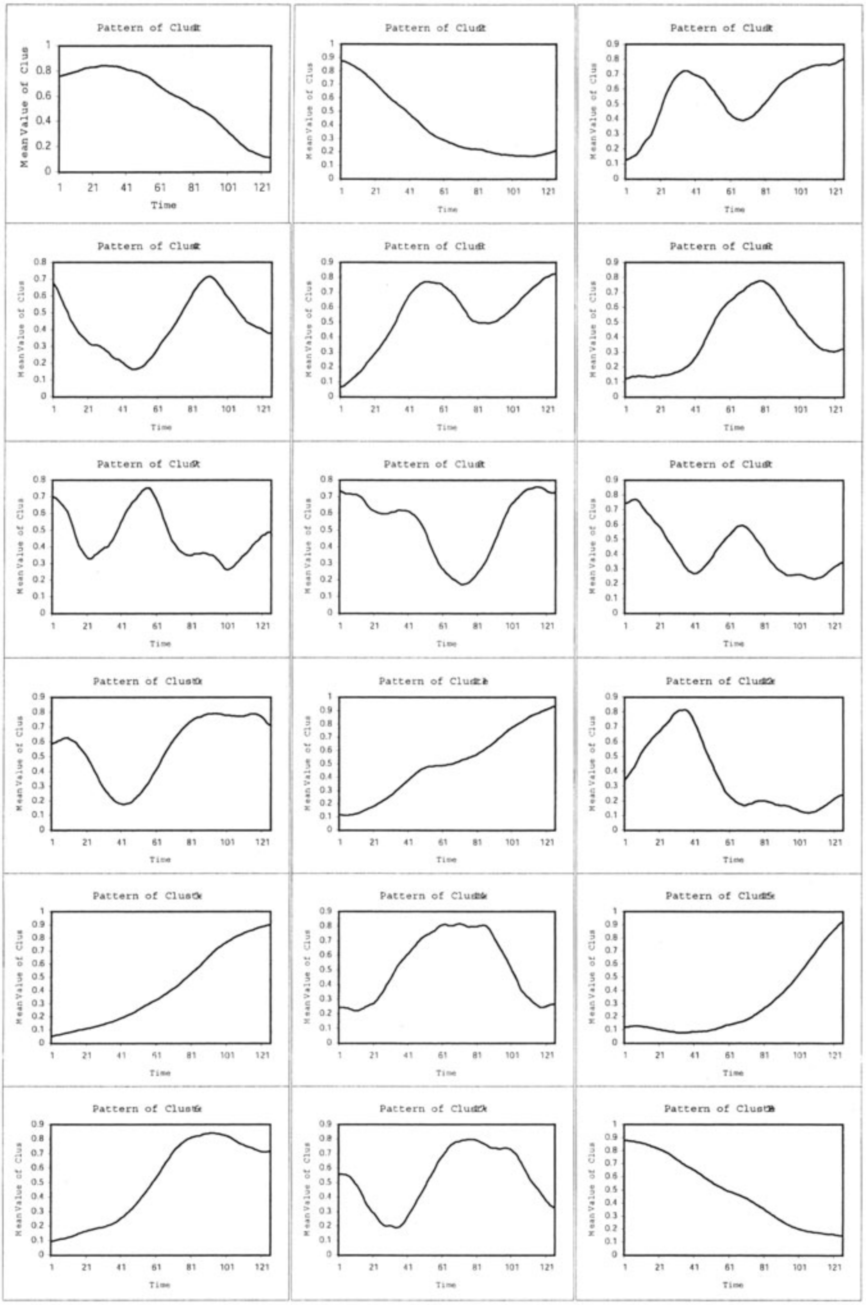

Fig. 8.2. Charts discovered by the 2-dimensional SOM: harts 1-18

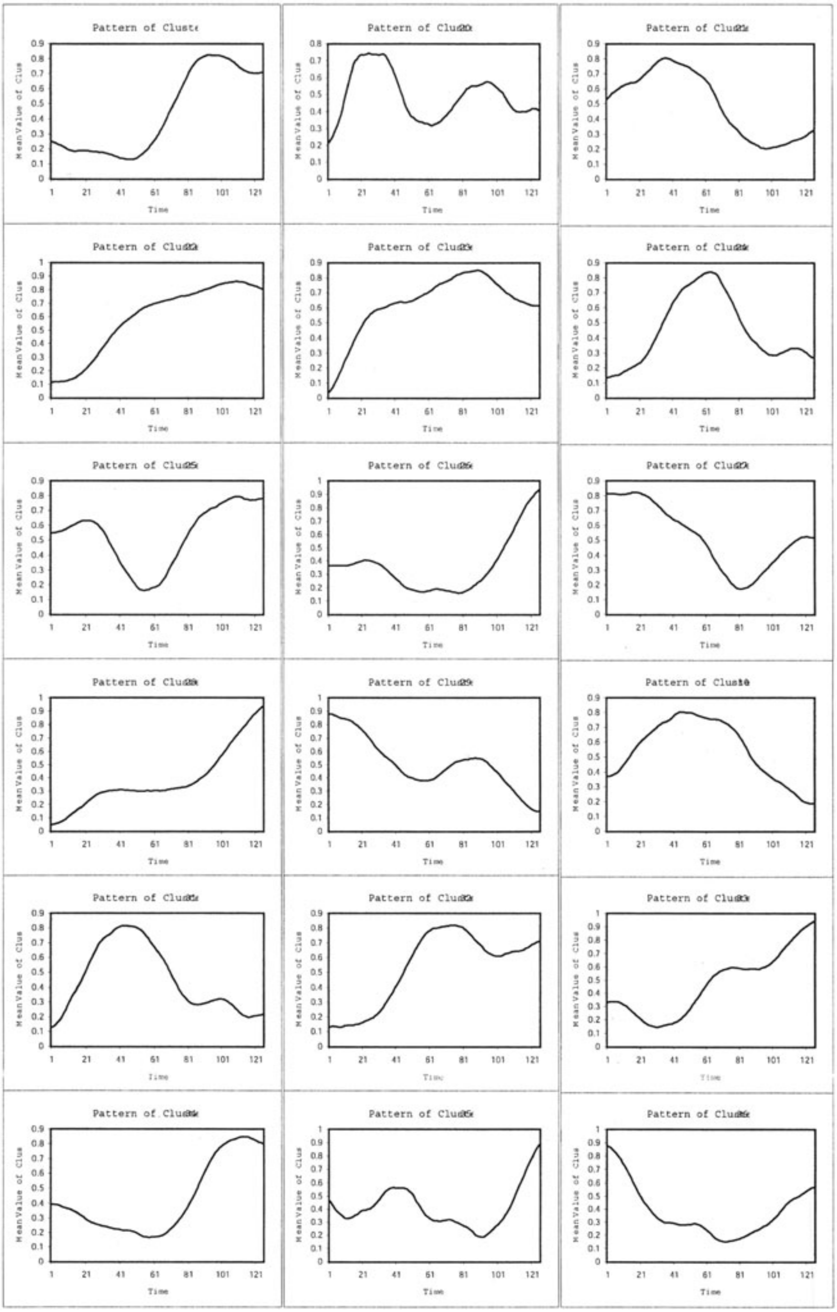

Fig. 8.3. Charts discovered by the 2-dimensional SOM: charts 19-36

surprisingly, all our charts behave quite *smoothly* as opposed to many of those used by chartists. This fundamental difference makes a direct comparison between these two types of charts not only infeasible but also meaningless. Nevertheless, some of our charts can roughly be given a name in chartists' eyes, such as *uptrends* (Charts 11, 13, 15), *downtrends* (Charts 1, 2, 18), *V-formations* (Charts 8, 24 and 25), *rounding bottom* (Chart 36), *rounding top* (Chart 30), *double tops* (Chart 20), *island reversal* (Chart 17).

8.4 Do SOM-Induced Charts Reveal Trading Signals?

So far, we only demonstrate how to use SOM to built charts. But, how do we know that these charts can indeed reveals trading signals rather than *just clusters*? To answer this question, we proposed the following test based on a *normalized equity curve.* First, we ask a simple question: *once a specific chart is observed, what are the stock returns in the following k days*? In other words, we are investigating the time series plot R_h,

$$R_{h,i} = \ln P_{t+h} - \ln P_t, \quad h = 1, 2, ..., k, \tag{8.6}$$

where t is the day which pattern (chart) i is observed ($i = 1, 2, ..., 36$). Call $R_{h,i}$ the *normalized equity curve* of the chart i. Then, we can make the following observation. Suppose that the patterns constructed by SOM are able to transmit trading signals, then what we expect from these equity curves are *systematic* movements, such as *monotone increasing* or *decreasing*, as the most ideal cases. They can be other types, but the bottom line is that they cannot be *erratic* or *random.* Finally, since each pattern appears many times in the whole series, instead of drawing a single equity curve, we are presenting an average of them, i.e., what we actually draw is the time series plot of $\bar{R}_{h,i}$:

$$\bar{R}_{h,i} = \frac{\sum_{j=1}^{m_i} R_{h,i:j}}{m_i}, \tag{8.7}$$

where j refers to the jth occurrence of chart i, and m_i is the total number of the occurrence of chart i.

Figs. 8.4 and 8.5 is the time series plot of these *representative* equity curves of the 36 charts. From these curves, most charts do seems to feature *strong* buy and sell signals. In 14 out of the 36 charts, the equity curve is monotonically increasing (Charts 2, 3, 4, 5,11, 12, 13, 16, 23, 26, 28, 32, 33, 35), featuring buying signals, and in 3 out of the 36, the equity curve is monotonically decreasing (Charts 1, 10, 18), featuring selling signals. For many others, even though the equity curve is not monotone, a trend to grow or decline is discernible. Therefore, we may roughly conclude from this initial analysis that the charts constructed by SOM are able to transmit buying and selling signals.

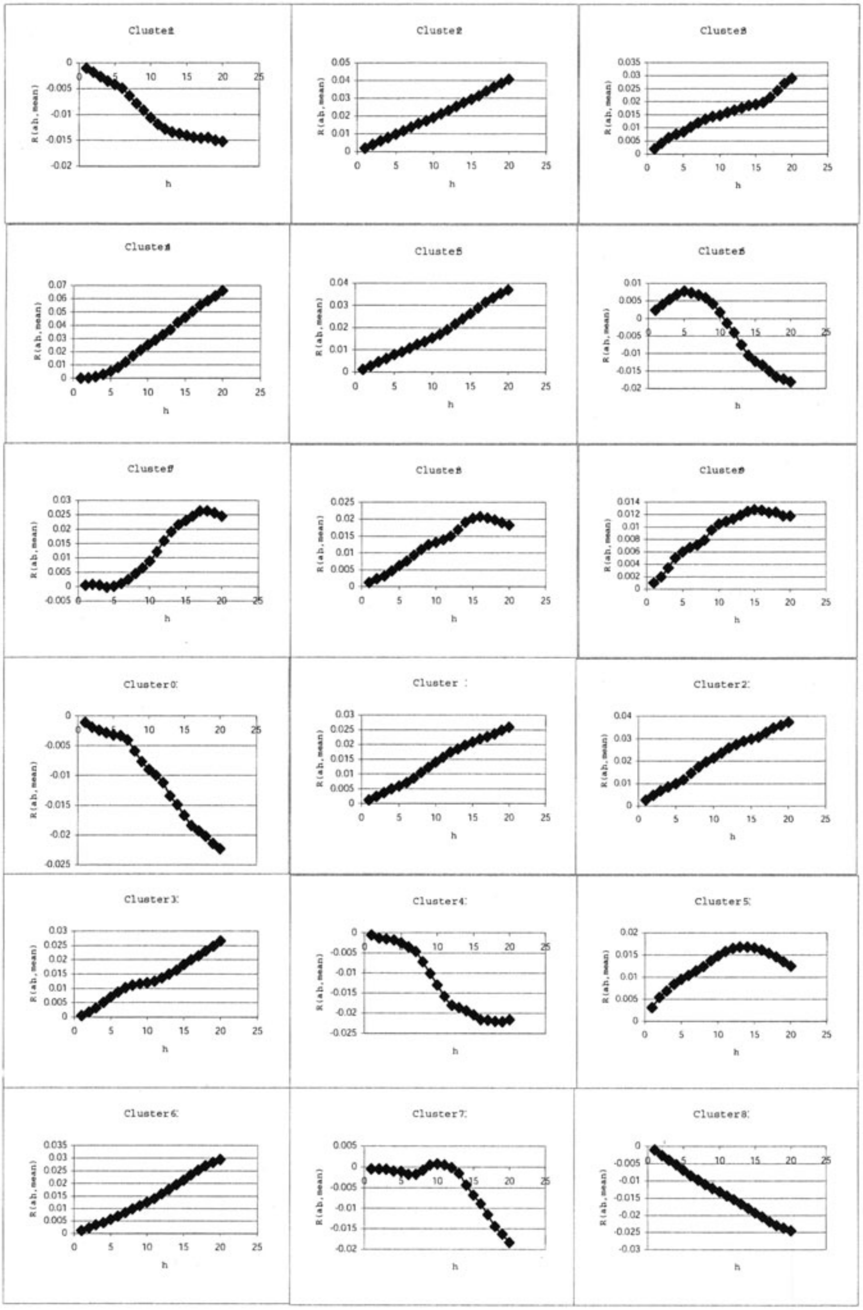

Fig. 8.4. The normalized equity curves of charts 1-18

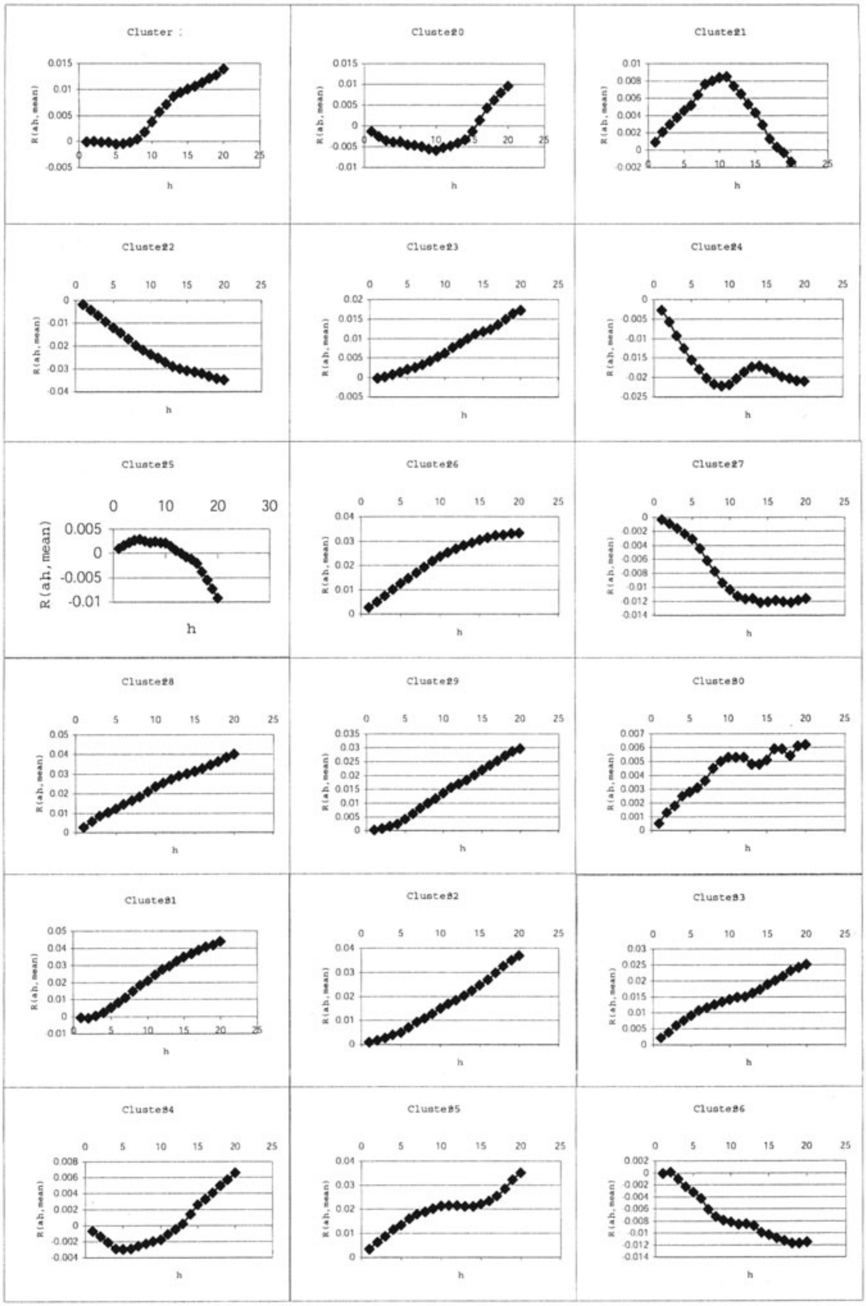

Fig. 8.5. The normalized equity curves of charts 19-36

8.5 SOM-Induced Trading Strategies

The previous analysis based on equity curves shows that the SOM-induced charts are not just revealing a bunch of similar patterns. The analysis indicates that many of these charts can be further used for helping us forecast the movement of stock price, at least, in the short term. In this section, we would like to go one step further, and ask: *if we actually use these charts to develop trading strategies, how well will they perform?* [3] Will they beat other well-known benchmarks, such as the *buy-and-hold* (**B&H** strategy)? But, how to trade with these charts? There are many possible ways to do so, the one considered in this paper is one of the simplest ones.

The first stage of this development is to associate with each chart a *buy* or *sell* signal. To do so, let us suppose that traders are interested in *day trading* only, and they are responsive to some patterns, say chart i ($i = 1, 2, ..., 36$) constructed by SOM. Usually, they will do two things when chart i appears, either *buy then sell* or *sell first then buy back later* (so-called *short selling*). Denote the first action, called the *buy* action, by **B**, and the second action, the *sell action*, by **S**. We can then calculate the rate of return for *one-time* appearance of Chart i. Now, let $R_{B,i}$ be the rate of return for Chart i under action B in the *whole* sample, and $R_{S,i}$ be that under action S. Then

$$R_{B,i} = \sum_{t=1}^{T} R_t I_{i,t} \tag{8.8}$$

where $I_{i,t}$ is an indicator function :

$$I_{i,t} = \begin{cases} 1, & \text{if Chart } i \text{ appears at time } t-1 \\ 0, & \text{otherwise.} \end{cases} \tag{8.9}$$

and $R_t = \ln P_t - \ln P_{t-1}$. The *mean* rate of return of trading on Chart i with the *buy* action can then be derived by taking the *simple average* as follows.

$$\bar{R}_{B,i} = \frac{R_{B,i}}{\sum_{t=1}^{T} I_{i,t}} \tag{8.10}$$

Similarly, we can have

$$R_{S,i} = -\sum_{t=1}^{T} R_t I_{1,t} \tag{8.11}$$

and

[3] To some extent, this inquiry is in line with the 2-stage process proposed by [8.8], namely, first, *visual segmentation* of the curve into primitive shapes of some portions, and second, a *knowledge-based correspondence* established between the time sequence of such primitive shapes found in the curve and specific investment decisions, including forecasts or trend assessments.

$$\bar{R}_{s,i} = \frac{R_{s,1}}{\sum_{t=1}^{T} I_{1,t}} \tag{8.12}$$

Table 8.2. Profitability of trading on the SOM-induced patterns: buy action

i	m_i	$\bar{R}_{B,i}$	A	i	m_i	$\bar{R}_{B,i}$	A
1	335	-0.101	S	19	195	-0.006	S
2	370	0.185	B	20	78	-0.132	S
3	123	0.203	B	21	181	0.089	B
4	88	0.003	B	22	207	-0.187	S
5	155	0.116	B	23	155	-0.016	S
6	150	0.236	B	24	154	-0.267	S
7	102	0.049	B	25	174	0.097	B
8	146	0.124	B	26	202	0.265	B
9	170	0.103	B	27	213	-0.032	S
10	136	-0.114	S	28	242	0.282	B
11	364	0.12	B	29	203	0.007	B
12	67	0.262	B	30	224	0.046	B
13	415	0.052	B	31	114	-0.073	S
14	164	-0.061	S	32	206	0.101	B
15	249	0.272	B	33	199	0.211	B
16	372	0.113	B	34	207	-0.072	S
17	108	-0.047	S	35	141	0.338	B
18	534	-0.112	S	36	166	-0.008	S

"i" denoted the ith chart, "m_i" indicates the frequencies of observing Chart i, and $\bar{R}_{B,i}$ is the mean rate of return for the buy action.

Table 8.2 summarizes these daily return statistics for each charts. We shall then use these return statistics described above to identify whether the appearance of a chart feature as a buy or a sell signal. For example, from Table 2, the first-day return for Chart 1 is negative (-0.00101), so it is recognized as a *sell* signal.[4] And it is 0.00185 for Chart 2, so Chart 2 features a *buy* signal. By this manner all charts are assigned to a **B** or **S** action, as shown in the fourth column of Table 8.2, and we can map the sequence of appearing charts to a sequence of "**B**" and "**S**" signals.

For example, in our case, the chart sequence starts with "34, 34, 34, 34, 34, 34, 34, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 25, 33, 33, 33, 33, 33, 33, 33, 33, 33, 33, 33, 33, 33, 33, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 18, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, ..." Then based on the description above, it can be mapped to the buy-or-sell sequence as below: "B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, S, B, B, B, B,

[4] As seen above, $\bar{R}_{S,i}$ is simply just the negative of $\bar{R}_{B,i}$. So, if the buy action earn a negative profit, then the sell action must earn a positive profit.

B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B,..." So, at the beginning, we start with the long position, i.e., to buy, and the trading strategy since then is constructed as follows.

Suppose at time $t-1$, the trader is in a long position (i.e., holding one unit of the stock on hand, **Holding** $=$ **1**). If at time t, there is a **B** signal, and the trader simply continue to **Hold** (i.e., do nothing, **Holding** $=$ **1**). However, if at time t, it is a **SELL** signal, then the trader with long position shall **SELL** her one unit of the stock kept in the previous period, and her holding is now 0 (**Holding** $=$**0**). In addition, she shall further **SELL SHORT** one more unit of stock and wait to buy it back later (i.e. waiting for the next coming buy signal). For the time being, she has a short position (**Holding** $=$ **-1**). Now, consider the other possibility. Suppose that at time $t-1$, the trader are in a short position (**Holding** $=$ **-1**), and at time t, there is a **B** signal. Then the trader will buy back the stock sold short earlier, and her position is therefore back to 0. But, in addition, she shall go further to buy one more unit of stock, and she end up with a long position (**Holding** $=$ **1**). On the other hand, if at t, it is a sell signal, then the trader will do nothing (**Holding** $=$ **-1**).

This automatic trading rule can be represented as a *mapping* G given below. Function G is a mapping from a Cartesian product space, which is a product space of the signal and position, to the position space,

$$G(x, y) = z, \tag{8.13}$$

where $x \in X = (B, S)$, and $y, z \in Y = (-1, 0, 1)$, with the following mapping: $G(B, 0) = 1, G(S, 0) = -1, G(B, 1) = 1, G(S, 1) = -1, G(B, -1) = 1, G(S, -1) = -1$. With this trading rule, we can then transform a **B-S** sequence into a sequence of 1, 0 and -1. Now, let the position at t as H_t (Clearly, $H_t = 1, 0$, or -1). Then (*accumulated*) trading profits at any point in time, π_t, can be calculated as follows,

$$\pi_1 = 0, \tag{8.14}$$

and

$$\pi_t = \begin{cases} \pi_{t-1}, \text{if} \quad H_t = H_{t-1}, \\ \pi_{t-1} + P_t(1 - c_1 - c_2) - P_{s(t)}(1 + c_1), \\ \text{if} \quad H_t = -1, H_{t-1} = 1, \\ \pi_{t-1} + P_{s(t)}(1 - c_1 - c_2) - P_t(1 + c_1), \\ \text{if} \quad H_t = 1, H_{t-1} = -1, \end{cases} \tag{8.15}$$

where

$$s(t) = \max_j \{j | H_j \neq H_t, H_{j-1} = H_t, 1 < j < t\}. \tag{8.16}$$

c_1 is the tax rate of each transaction, and c_2 the tax rate of securities exchange income. In the case of Taiwan, $c_1 = 0.001425$, and $c_2 = 0.003$.[5]

[5] This way to evaluate the accumulated profits is very standard and has been extensively used by practitioners. Also see [8.1].

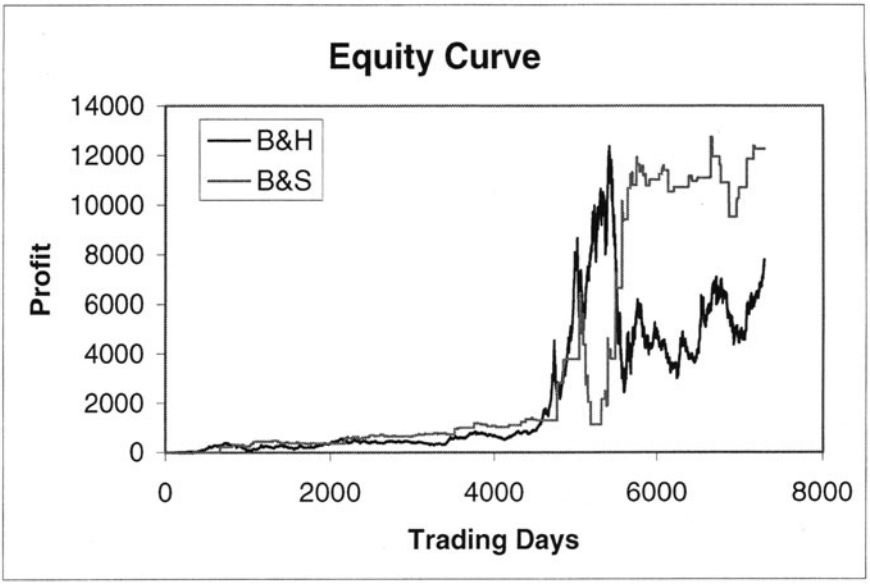

Fig. 8.6. The equity curves of the buy-and-hold strategy and the SOM-based trading strategy

Time series plot for π_t of the SOM-based trading strategy has been overlaid on that of the benchmark **B&H** strategy in Fig. 8.6. From these two plots, there are *roughly* two noticeable crossing points, one appears around the 5000th trading day, and the other appears near the 5500th day. These two crossing points suggest that neither trading strategy can dominate the other. However, the period that **B&H** outperforms **SOM** is much shorter than the other, and it occurs only between the first crossing point and the second one, which is about 500 days. For the rest of time, in particular, after the second crossing point, **SOM** dominates **B&H**. And SOM keeps the lead till end with the accumulated profits NT$ 12,258, which is 60% higher than **B&H** at that day.

8.6 Concluding Remarks

The ability to recognize patterns is an essential aspect of human intelligence. Herbert Simon, who won a Nobel Prize in economics in 1978, consider pattern recognition critical and advocates the need to pay much more explicit attention to teaching pattern recognition. Chartists appear to be good at doing pattern recognition for many decades, yet little academic research has been devoted to a systematic study of this kind of activities. [6]Lo, Mamaysky

[6] Apart from this paper, to our best knowledge, the only other attempt is [8.7].

and Wang (1999). On the contrary, sometimes it was treated as nothing more than the astrology, and hardly considered as a science. Using Self-organizing maps, this paper proposes a systematic and automatic approach to *charting*. To some extent, what SOM does is to *simulate* human intelligence in finding or creating patterns that summarize and store useful aspects of our perceptions.

In addition to using SOM simulate this cognitive process, our initial analysis also evidences why charting can be essential in technical analysis. Applying SOM to a 25-year time series data of Taiwan stock index, we have two interesting results. First, charts constructed by SOM can reveals profitable signals. Second, a simple trading strategy built upon charts can beat the buy-and-hold strategy. Moreover, the efforts to get these results is quite limited; we do not tried lots of variants of running SOM, which is certainly an interesting direction for the further study.

Acknowledgements

Research support from NSC grant No. 90-2415-H-004-018 is gratefully acknowledged. The chapter was refereed by Chueh-Yung Tsao. The authors are grateful to his corrections of some mistakes made in the earlier version of the paper.

References

8.1 Chen, S. -H., Chen, C. -F. (1998): Can We Believe that Genetic Algorithms Would Help without Actually Seeing Them Work in Financial Data Mining? Part 2, Empirical Tests. In: Xu, L., Chan, L. W., King, I., Fu, A. (Eds.), Intelligent Data Engineering and Learning: Perspectives on Financial Engineering and Data Mining, Springer
8.2 DeMark, T. R. (1994): The New Science of Technical Analysis. Wiley
8.3 Haykin, S. (1994): Neural Networks: A Comprehensive Foundation. MacMillan
8.4 Huang, Z. (1997): A Fast Clustering Algorithm to Cluster Very Large Categorical Data Sets in Data Mining. First Asia Pacific Conference on Knowledge Discovery and Data Mining. World Scientific
8.5 Kohonen, T. (1982): Self-organized Foundation of Topologically Correct Feature Maps. Biological Cybernetics, Vol. 43, 59–69
8.6 Kohonen, T. (1994): Self-Organizing Maps. Springer
8.7 Lo, A., Mamaysky H., Wang, J. (2000): Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation. Journal of Finance, Vol. 55, No. 4, 1705–1770
8.8 Pau, L. F. (1991): Technical Analysis for Portfolio Trading by Syntactic Pattern Recognition. Journal of Economic Dynamics and Control, 715–730